Published in final edited form as:

Conf Proc IEEE Eng Med Biol Soc. 2010 ; 2010: 4312–4315. doi:10.1109/IEMBS.2010.5626204.

Automatic Detection of Dining Plates for Image-Based Dietary Evaluation

Jie Nie,

Department of Computer Science, Ocean University of China, Qingdao, China

Departments of Neurosurgery / Electrical Engineering / Bioengineering / Psychiatry, University of Pittsburgh

Zhiqiang Wei,

Department of Computer Science, Ocean University of China, Qingdao, China

Wenyan Jia,

Departments of Neurosurgery / Electrical Engineering / Bioengineering / Psychiatry, University of Pittsburgh

Lu Li,

Departments of Neurosurgery / Electrical Engineering / Bioengineering / Psychiatry, University of Pittsburgh

John D. Fernstrom,

Departments of Neurosurgery / Electrical Engineering / Bioengineering / Psychiatry, University of Pittsburgh

Robert J. Sclabassi, and

Departments of Neurosurgery / Electrical Engineering / Bioengineering / Psychiatry, University of Pittsburgh

Mingui Sun

Departments of Neurosurgery / Electrical Engineering / Bioengineering / Psychiatry, University of Pittsburgh

Mingui Sun: drsun@pitt.edu

Abstract

An automatic detector that finds circular dining plates in chronically recorded images or videos is reported for the study of food intake and obesity. We first detect edges from input images. After a number of processing steps that convert edges into curves, arc filtering and grouping algorithms are applied. Then, convex hulls are identified and the ones that fit the description of ellipses corresponding to dining plates are determined. Our experiments using real-world images indicate that this detector is highly reliable and robust even when the input images contain complex background scenes and the dining plates are severely occluded.

I. Introduction

As obesity spreads rapidly around the world, it is essential to know what, how, and how much people eat in real-life. However, existing methods for dietary evaluation are usually

based on self-reporting, which is affected by limited human memory and a number of psychological factors. As a result, the amount of food intake is often underreported. In order to overcome this problem, we have developed a wearable device which images the region in front of the wearer continuously during the wake hours. The images captured by our device are used to obtain dietary information objectively [1, 8]. Because the chronically recorded image dataset is large, automatic detection of eating events from this dataset is highly desirable in order for the dietitian to estimate food portion sizes and evaluate calories/nutrients efficiently. Because direct recognition of human food from images by computational means is extremely difficult, we turn to automatic detection of dining plates which mark a great portion of eating events. Plate detection is also the first, but a critical, step for further automatic food recognition and measurement because this detection provides well-defined regions to apply computational algorithms.

In this work, we only consider circular plates. It has been shown that, a circular plate always has an elliptical contour in images. Therefore, our task of plate detection reduces to the detection of certain forms of ellipses. Detection of elliptical shapes has been studied extensively in the field of computer vision. The existing methods can be roughly categorized into two classes: Hough based methods [4, 7] and elliptical model fitting methods [2]. The principal idea in Hough based methods are the use of pixels in the spatial domain to “vote” in a parameter space where the detection is determined by the most votes. However, in practical cases, using the Standard Hough transform (SHT) to detect ellipse has been a challenge due to the involvement of high computational complexity [4]. Because of this problem, many improved versions of the Hough transform have been proposed to reduce storage and computational cost, such as by randomization in the voting scheme using the Randomized Hough Transform (RHT) [4], and utilization of geometrical properties of the ellipse [6, 7]. Despite these, in our experiments, Hough transform based methods did not perform well in dining plate detection because these plates were often occluded either by food items or by nearby objects on the dining table. Therefore, many geometrical properties that the algorithms depend on are difficult to extract. In real-world images, certain non-target objects may increase votes to the incorrect parameters, while certain occlusions decrease votes to the correct parameters. As a result, the parameters obtained from Hough based methods are often unreliable. The methods based on elliptical model fitting detect ellipses by minimizing a certain distance measure between data points and the elliptical model. For example, the Direct Least Square (DLS) method [2] utilizes a least square elliptical fit to a given set of data points. The fitting result is reliable when the data points correspond to the whole or a partial ellipse. However, in real-world images, edges of clutters are often connected with dining plate edges, and it is difficult to determine useful data from clutters.

The splitting-merging algorithm reported in [3] describes an effective approach to the connection problem between outliers and target objects. This algorithm splits edges into partitions and then selectively merges them to find correct curve combinations by evaluating the quality of ellipses fitted by DLS. However, this algorithm only breaks the connections between outliers and target objects to a certain degree, and it does not remove outliers themselves. Typically, there are numerous curves to fit, which demands expensive computation in the merging phase of the algorithm, even when the image content is moderate. In addition, the outliers in images often produce numerous erroneous fit with faulty high scores, which increases the false positive error.

In this paper, we proposed a dining plate detection method by splitting, filtering and grouping the edges of objects. The new approach can detect severely occluded dining plates in a cluttered image background. We also present new concepts of arc filtering and grouping (highlighted in Fig.1). We probe edges of dining plates based on statistical properties of arc

lengths to remove outliers in the filtering phase of the algorithm. As a result, both the computational efficiency and detection performance are improved significantly.

II. Methods

Fig.1 depicts the framework of our dining plate detection method. In order to increase computational efficiency, we use edge map as the input to our algorithm. Since this binary map (or image) can be obtained from any high-performance edge detection algorithms which are widely available, we will not discuss the details.

A. Curve Drawing and Splitting

Since it is difficult to computationally pick edge pixels of a plate directly from the edge map, a line drawing algorithm described in [5] is adopt to restore connectivity of edge pixels. As a result, a set of curved lines (or curves) are formed from the discrete pixels in the edge map. The general form of the curve equation is given by $y = f(x)$. In order to break erroneous connections between close but different objects, we split all curves at three types of points (x_c, y_c) : (1) $f(x_c)$ is not differentiable (the tangent line is vertical), (2) $f'(x_c) = 0$ (the tangent line is horizontal), and (3) $f''(x_c) = 0$ (an inflection point). We name the curve elements after splitting as “arcs”.

B. Optimal Arc Filtering

When compared to edges of dining plates, the edges of most clutters in the input image are usually more irregular. As a result, the average length of the clutter arc is generally shorter than the plate arc. Based on this observation, we propose an optimal filter using the statistical properties of this length difference.

In order to obtain the desired statistical properties, 220 real life images containing dishes of food were acquired by a self-constructed wearable device [1]. We used the curve splitting criteria just described and constructed histograms of the plate arc lengths and clutter arc lengths. We found that, in either case, the normalized histograms can be modeled closely by

the Rayleigh distribution, given by $f(x; \sigma) = \frac{x}{\sigma^2} \exp(-\frac{x^2}{2\sigma^2})$ with $\sigma = \sigma_{clutter}$ for the clutter arcs and $\sigma = \sigma_{plate}$ for the plate arcs.

We define the optimal threshold of filtering x_0 by:

$$x_0 = \underset{x>0}{\operatorname{argmin}}(Pe) = \sqrt{2 \ln \left(\frac{P_{clutter}}{P_{plate}} \times \frac{\sigma_{plate}^2}{\sigma_{clutter}^2} \right) / \left(\frac{1}{\sigma_{clutter}^2} - \frac{1}{\sigma_{plate}^2} \right)} \quad (1)$$

with:

$$Pe = P_{clutter|plate} \times P_{plate} + P_{plate|clutter} \times P_{clutter} \quad P_{clutter|plate} = 1 - \int_x f(x; \sigma_{plate}) dx \quad P_{plate|clutter} = \int_x f(x; \sigma_{clutter}) dx$$

and $P_{clutter}$ and P_{plate} are obtained from the Rayleigh density model. During filtering, only arcs with lengths greater than x_0 are kept. This procedure removes a vast majority of noise arcs, as indicated by our experiment (to be described).

C. Curve Grouping

The previously described splitting and filtering procedures break erroneous connections between objects and remove most clutter arcs. We then implement the following grouping procedure to reconnect arcs that belong to the same object.

One of the important features of the ellipse is its convexity. It is well-known that a convex object is always on the same side of the tangent line passing through any of its boundary point. We illustrate this property in Fig. 3 where every arc A_i divides image plane I into two parts B_i and $\bar{B}_i = I - B_i$, where B_i is the region defined by the tangent line of the center point of arc A_i directed to the concave side of the arc. (shown in Fig. 3 as red arrows). As an example, the shaded region in Fig. 3 indicates B_1 of arc A_1 . We collect all possible convex shapes in image I in $H = \{h_q\}$, $q = 1, \dots, m$, $h_q = \{A_i | i = n_{q1}, \dots, n_{qm}\}$ with m being the number of possible convex shapes. The criteria utilized are:

$$\text{if } A_i, A_j \in h_q, \text{ then } A_i \in B_j \text{ and } A_j \in B_i, i, j = 1, \dots, n; q = 1, \dots, m$$

An upper triangular conjunction matrix C can be formed to represent this relationship between every two arcs:

$$C = \begin{bmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & c_{nm} \end{bmatrix}, c_{ij} = \begin{cases} 1, & \text{if } A_i \in B_j \text{ and } A_j \in B_i \\ 0, & \text{otherwise} \end{cases}$$

According to Graph theory, finding H is equivalent to finding all cliques from an undirected graph using matrix C .

D. Ellipse Identification

After grouping arcs into convex shapes, we determine whether any of these convex shapes are ellipses. Each convex group of arcs is fit using the DLS method. These ellipses are rated by a quality measure defined as:

$$Q(e_i) = \frac{\text{Num}(\text{Pixel}(e_i) \cap \text{Pixel}(h_i))}{\text{Num}(\text{Pixel}(e_i))}, i = 1, \dots, m \quad (2)$$

Where h_i is the i^{th} convex shape, e_i is the template ellipse fit by arcs in h_i , $\text{Pixel}(X)$ returns the pixel set of X (for $X = h_i$ it returns pixels consisting of arcs in the convex shape; whereas for $X = e_i$ it returns pixels on the entire boundary of the template ellipse), $\text{Num}(X)$ returns the number of pixels of X , and $\cap$ is an intersection operator. When calculating the intersecting pixels of the template ellipse and arcs, we also include the four neighborhood pixels of every pixel of each arc. This inclusion avoids underestimation of Q due to slight deviations of arc locations in images.

The ellipse quality Q provides the ratio of the number of arcs pixels that overlap with the boundary of the template ellipse and the number of pixels of the entire template ellipse. Each plate candidate fitted by one convex hull is determined to be a real plate only if its quality value is greater than a threshold α and the ellipse has appropriate size and orientation in the image. The threshold and the size/orientation criteria are case dependent thus must be determined empirically.

III. EXPERIMENT

The new dining plate detection method was tested using the images acquired by our wearable device in real world [1]. Without active control of the camera, the images, in a screen size of 800×600, are noisier and more cluttered than the images obtained by usual means. Our images contain approximately 98% clutter arcs (as defined previously) and 30% occlusion. Here the occlusion is defined as the ratio between the visible edge pixels of a plate and the edge pixels of the entire plate, regardless of visibility. For our images, the Hough based ellipse detectors, which rely heavily on the geometrical properties of the visible elliptic boundaries, were ineffective, as shown by our experiment. In addition, due the plates always contain food, the visible parts of the plate edges in our case were almost always discontinuous, creating more difficulties.

We pre-processed our input images using the blind deconvolution algorithm for deblurring, and histogram equalization for reducing illumination inhomogeneity and achieving a higher contrast. Then, the previously described computational procedures were applied. We performed two experiments to valid our method.

In the first experiment, we tested ellipse fitting accuracy using 20 images; each of these images contained a single dish. We evaluated the ellipse fitting error with respect to each elliptic parameter by dividing the absolute difference between the true and detected values by the true value, where the true value was obtained by marking the ellipse boundary pixel manually and then fitting the marked pixels by DLS. Our results were summarized in Table 1. The parameters listed in the first row of the tale are the coordinates of ellipse center (x , y), the lengths of the short and long axes α and β , respectively, and the orientation θ of long axis β with respect to the x -axis. Table 1 shows that our algorithm achieves high detection accuracy, demonstrating that the procedures of removing clutter arcs and grouping plate arcs are effective.

In the second experiment, the performance of the ellipse identification procedure was evaluated. 35 real-life images with occluded dining plates and an equal number of images without plates were utilized in this experiment. The threshold α for ellipse identification was experimentally selected in arrange between 0.1 and 1.0 with an increment of 0.05. Fig. 4 shows the Receiver Operating Characteristic (ROC) curve in dinning plate identification. For example, the black rectangle in Fig.4 (see the amplified view) corresponds to $\alpha = 0.3$ which indicates that, on average, a plate occluded less than 70% can be identified with a true positive ratio of 97.14% and a false positive ratio of 2.86%. These values give rise to a false negative ratio of the same 2.86%.

We illustrate the results of dining plate detection in Fig.5. The detecting and fitting results are superimposed on the original images by red ellipses (Figs. 5a and 5d). It can be observed that the dining plate in Fig. 5a was severely occluded by food items from the inside of the plate. Despite this, our method broke plate boundaries at both critical points and inflections correctly (shown in Fig.5(b)(e)), discarded clutter arcs by length filtering (shown in Fig.5(c)(f)), and finally rebuilt connections of arcs by the grouping procedure. Similarly, it can be observed from Figs. 5d through 5f that our method was also able to detect multiple occluded plates in a cluttered image background.

IV. CONCLUSION

In this paper, we have presented an ellipse detection method by splitting, filtering and grouping edges. This method is reliable in detecting extremely occluded ellipses in noisy images with complex background, as shown by our experimental results.

Our study indicates that, by breaking low level edges into elementary arcs, performing a certain classification task (e.g., arc filtering), then re-grouping the filtered arcs to form features in a higher level, we can better understand image content and object shapes. Thus, the method presented in this paper provides a useful approach to object detection in general, where the objects are not limited to ellipses.

As an important application, our method can be used to detect food intake events automatically by indentifying dining plates in chronically recorded video acquired by a wearable device. This method also provides a platform for automatic food recognition since it defines the regions to apply food recognition algorithms.

We note here that our method only make use of the information provided by edges. While this approach is computationally efficient, it limits the understanding of image content in other domains, such as color and texture. However, information in these domains can be incorporated into the identification phase of the method to improve performance at the cost of a certain loss of efficiency.

Acknowledgments

This work was supported by National Institutes of Health Grant No. U01 HL91736 of U.S.A, the Key International Cooperation Grant funded by Science and Technology Ministry of China No. 2008DFA11030, Program for New Century Excellent Talents in University No. 2006NCET-06-0600, Doctoral Fund for new teacher of Ministry of Education of China No. 20090132120013 and National High Technology Research and Development Program No. 2009AA12Z330.

References

1. Sun M, Fernstrom JD, Jia W, Hackworth SA, Yao N, Li Y. A wearable electronic system for objective dietary assessment. *American Dietetic Association*. 2010 Jan.:45–47.
2. Fitzgibbon AW, Pilu M, Fisher RB. Direct least squares fitting of ellipse. *IEEE Trans. PAMI*. 1999:476–480.
3. Chia, Alex YS.; Rajan, D.; Leung, Maylor KH. A split and merged based ellipse detector; *Proc. IEEE Int. Conf. Image Processing*; 2008. p. 3212–3215.
4. Xu L, Oja E, Kultanen P. A new curve detection method: Randomized hough transform. *Pattern Recognition Letters*. 1990:331–338.
5. Leung MK, Yang Y. Dynamic two-strip algorithm in curve fitter. *Pattern Recognition*. 1990:69–79.
6. Ho C, Chen L. A fast ellipse/circle detector using geometric symmetry. *Pattern Recognition*. 1995:117–124.
7. Chia, A.; Leung, M.; Eng, H.; Rahardja, S. Ellipse Detection with hough transform in one dimensional parametric space; *Proc. IEEE Int. Conf. Image Processing*; 2007. p. 333–336.
8. Nie, J.; Fernstrom, JD.; Scabassi, RJ.; Wei, Z.; Li, L.; Zhang, W.; Jia, W.; Mao, Z.; Sun, M. Automatic detection of dining plates in digital video. *Proc. 36th Northeast Bioengineering Conference*; New York. 2010.

Fig. 1.
Block Diagram of dining plate detection

Fig. 2.
Probability densities of clutter arcs (narrow peak) and dining plate arcs (wide peak).

Fig. 3.
Grouping Arcs by convexity

Fig. 4.
ROC curve of dinning plates detection

Fig. 5.
 (a)(d) Original images superimposed by red ellipse detected by our method. (b)(e) Edge images after splitting. (c)(f) Edge images after filtering.
 Note: this figure is best viewed in color.

TABLE I

Percentage Error of Parameters

	x	y	α	β	θ
<i>Mean</i>	0.96%	1.07%	3.68%	2.02%	2.38%
<i>Std</i>	0.78%	1.38%	1.92%	1.42%	1.54%